



# TRANSFORMER-BASED JOINT-STATION SEISMOLOGICAL ANALYSIS APPLIED TO FOCAL-MECHANISM DETERMINATION

ETH Zürich

Xiaohan Song<sup>1</sup>, Men-Andrin Meier<sup>2</sup>, William L. Ellsworth<sup>1</sup>, Gregory C. Beroza<sup>1</sup>

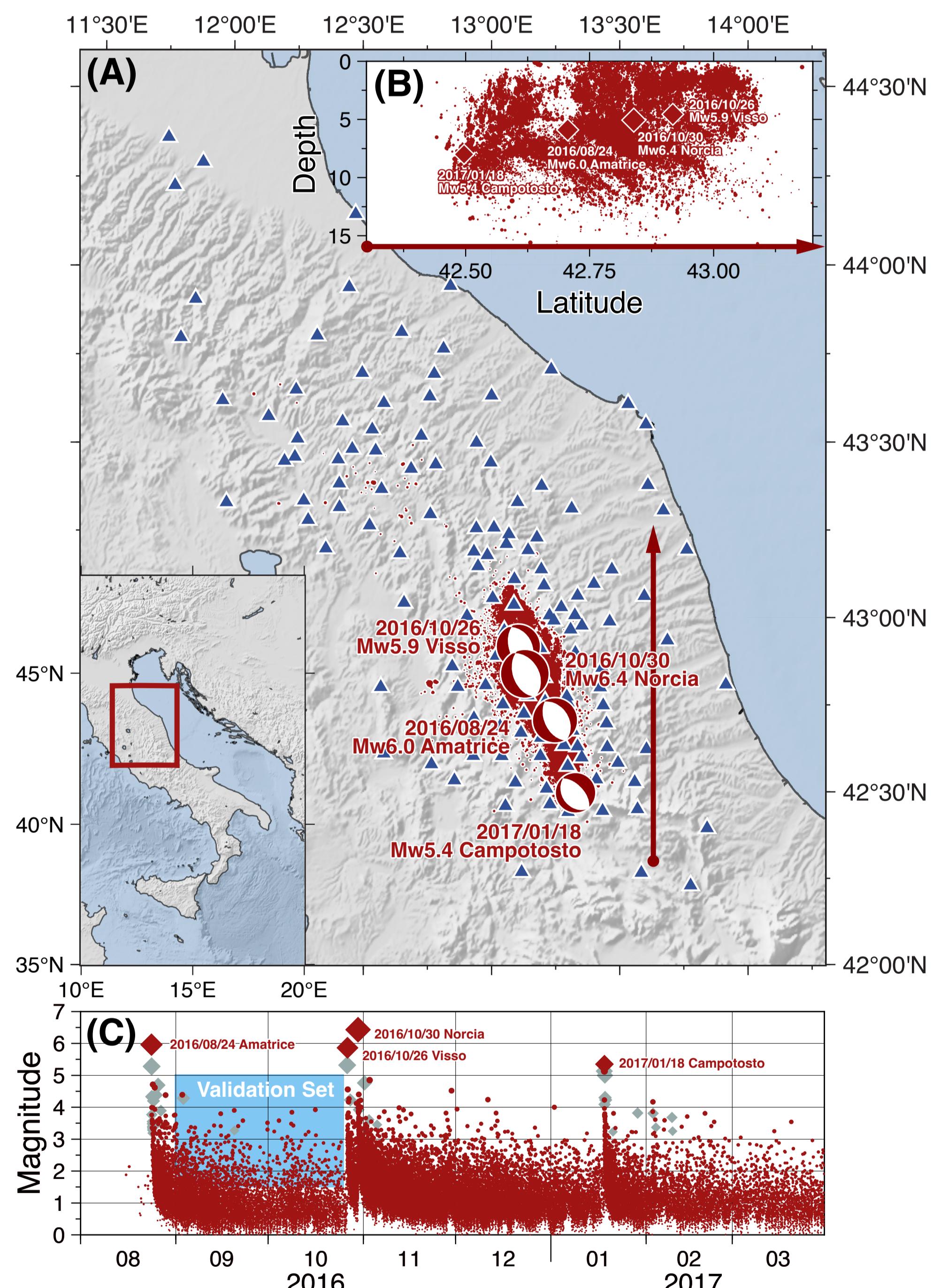
<sup>1</sup>. Department of Geophysics, Stanford University, CA, 94305    <sup>2</sup>. Institut für Geophysik, ETH Zürich, 8092 Zürich, Switzerland

## ① MOTIVATION

Deep learning models have been widely applied in processing seismic signals. But when it comes to combining phase information (e.g., arrivals, polarities, amplitudes) across a station network, most workflows rely on either physics-informed optimization (e.g., grid search), or neural networks trained on fixed station arrays (which limits their application to specific areas).

We propose that by treating every station's phase information as a "token vector" and appending them with the station locations/metadata as "positional encodings", we can train a transformer-based neural network capable of processing phase across any distributed station arrays.

We aim to build a transformer-based focal mechanism (FM) determination model and validate it on an FM dataset from the 2016 Italian Amatrice sequence [1].



## REFERENCES

- [1] Meier et al., A deep catalogue of 56k focal mechanisms for the.... EGU, 2023.
- [2] Xi et al., PyFk: A Fast MPI and CUDA Accelerated.... AGU, 2021.
- [3] Reasenberg and Oppenheimer, FPFIT, FPPILOT and FPPA.... U.S. Geological Survey, reports 1985.
- [4] Hardebeck and Shearer, Using S/P Amplitude Ratios to.... BSSA, 93(6), 2434–2444, 2003.
- [5] Skoumal et al., SKHASH: A Python Package for Computing.... SRL, 95(4), 2519-2526, 2024.

## ② FOCONET MODEL STRUCTURE

Our transformer-based model, named FocoNet, whose key part consists of 7 layers of multihead self-attention, directly solves for the focal mechanism based on the relative locations, first-motion polarities, S/P amplitude ratios, and SNRs from a set of stations. Ideally, we assume that the polarities have been determined by a pre-existing neural network (e.g., Meier et al.'s CNNs [1]), and we know the source locations & phase arrivals.

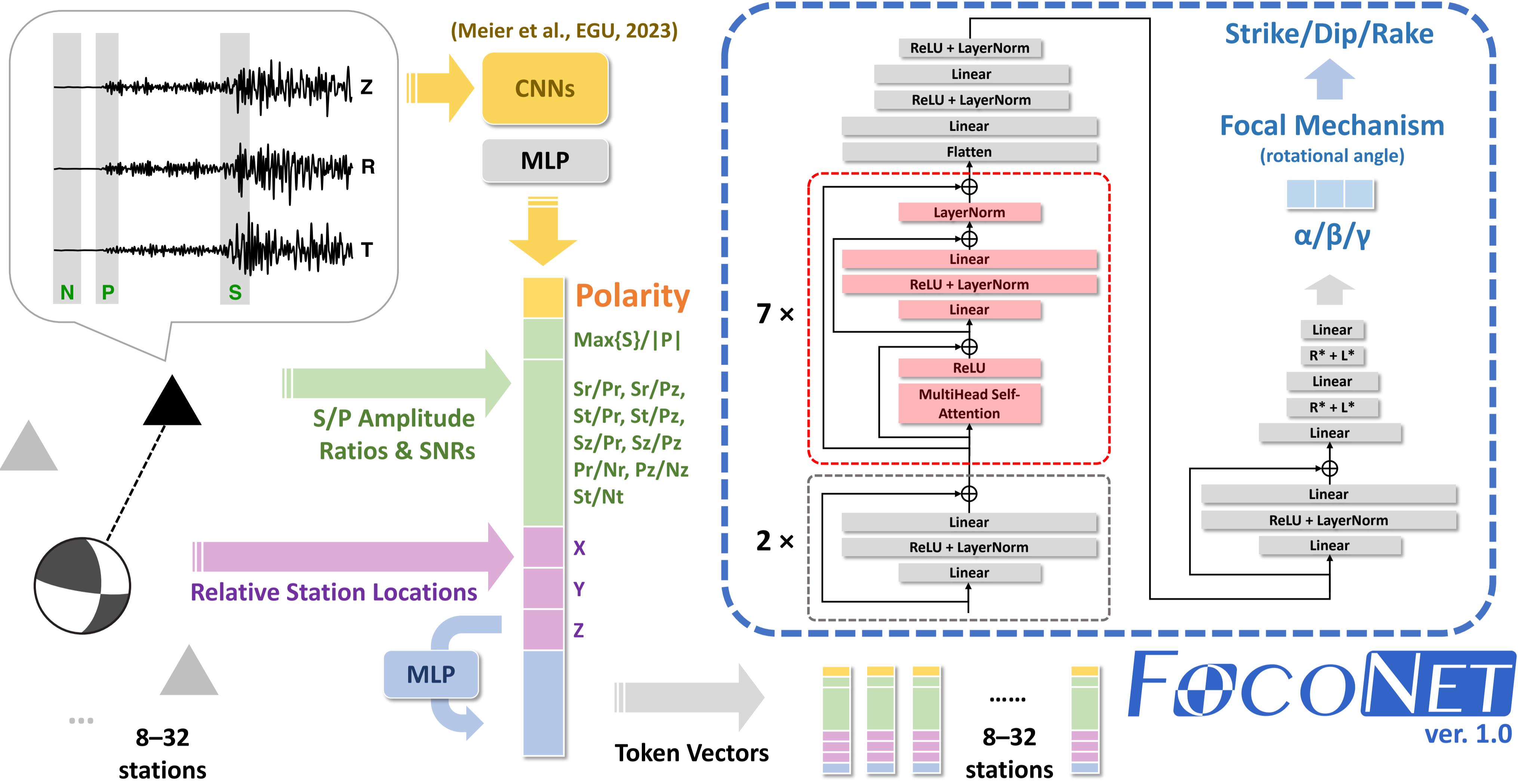


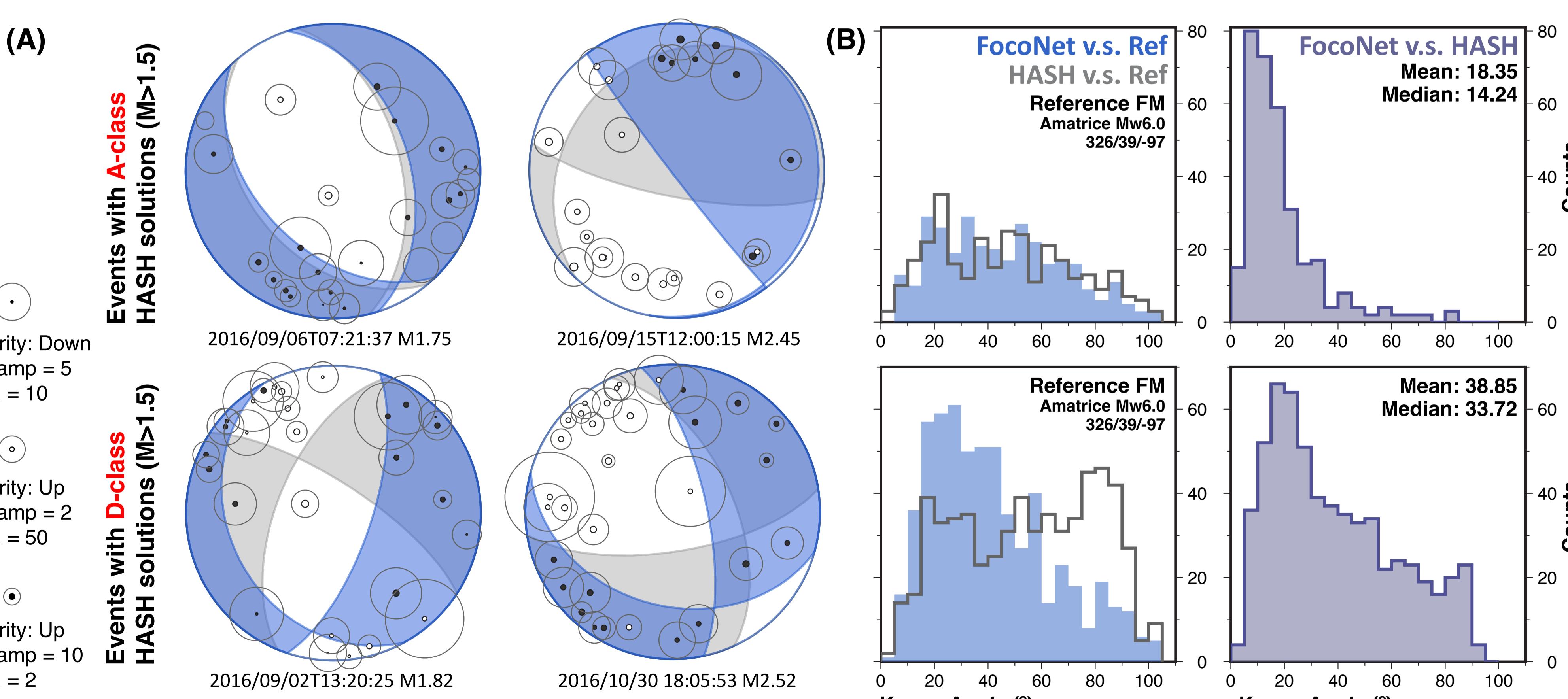
Figure 2: The workflow to solve for a focal mechanism and the model structure of our FOCONET

## ④ VALIDATION ON THE AMATRICE SEQUENCE

We validate FocoNet on 300+ Amatrice sequence events with A-class HASH solutions and 600+ events with D-class solutions. For these A-class events, FocoNet predictions achieve only 18° of Kagan angle difference compared to their high-quality HASH solutions. For these D-class solutions, FocoNet predictions effectively diminish the "random distribution peak (70°)" in the histograms of Kagan angles compared to the reference FM.

The transformer-based FocoNet can efficiently combine different kinds of phase information across a randomly distributed station array, and give far more accurate solutions than traditional grid search methods can do.

Both the enriched types of phase information (HASH standard S/P ratios, S/P ratios from different channels, and SNRs) and the increased ability to address uncertainty make FocoNet a superior tool than the traditional methods.



## ③ TRAINING AND TEST

To train FocoNet, we create 500,000 randomly rotated focal mechanisms. For each event, we randomly distribute 8–32 stations within 120-km epicenter distance. We generate synthetic waveforms through PyFk [2]. We randomly added noise to the synthetics to mimic the Amatrice real waveform as our training and test sets. We use the average Kagan angle between our predictions and the labels as the Loss function.

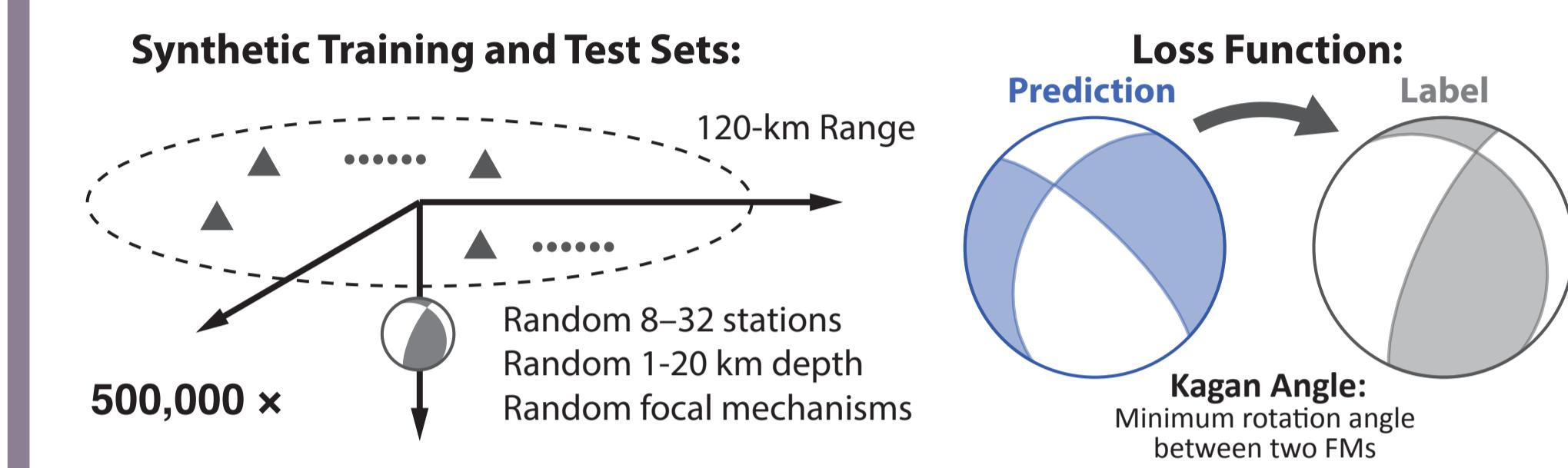


Figure 3: Our training/test set and loss function

We test FocoNet on different synthetic test cases and compare the predictions with those from traditional methods (e.g., FPFIT [3] and HASH [4, 5]). FocoNet consistently gives the most accurate predictions in all of the test cases.

Methods Test Sets	Kagan Angle (°)			
	FPFIT	SKHASH (excluding S/P)	SKHASH (including S/P)	FOCONET
12 stations 50 km	32.38 26.01	35.91 30.82	29.43 23.09	<b>18.08 14.16</b>
24 stations	21.79 15.87	20.25 16.44	17.30 13.26	<b>14.45 9.28</b>
32 stations	18.67 13.99	17.11 12.76	14.45 11.70	<b>9.98 8.59</b>
24 stations 90° Azimuth Gap	24.09 17.83	23.39 18.29	19.28 14.41	<b>14.00 11.45</b>
24 stations 180° Azimuth Gap	29.76 24.06	28.14 23.34	22.72 17.12	<b>17.40 14.17</b>
24 stations 270° Azimuth Gap	43.97 40.46	43.50 42.05	34.29 27.74	<b>30.43 24.62</b>
12 stations 10% Polarities Flipped	40.61 33.98	40.89 35.24	33.25 25.06	<b>23.04 17.66</b>
24 stations 10% Polarities Flipped	30.07 23.16	25.77 20.02	21.73 15.77	<b>14.60 12.03</b>
32 stations 10% Polarities Flipped	26.06 21.02	22.12 16.50	18.08 13.01	<b>13.08 10.31</b>

Table 1: The comparison between FOCONET and traditional methods on synthetic

## CONTACT

Xiaohan Song:  
xhsong@stanford.edu

